

Review Summarization using Feature-based Sentence Categorization

Presented by : Tan Zhiyu

Co-author : Teh Xin Xi

Wong Kean Yi

Gan Keng Hoon

What

Why





Outline

- Objective
- Proposed Methodology
- Proposed Framework
- Limitation
- Conclusion

Objective

- Entity Detection
- Sentiment Extraction
- Prediction Rating



Proposed Methodology

- 1) Pre-processing
- 2) Entity Detection
- 3) Post-processing

1) Pre-processing

- Boundary Detection
- Sentence Tokenization

The room was clean and very nicely decorated.



| | | | | | | | | |
|------------|-------------|------------|--------------|------------|-------------|---------------|------------------|----------|
| <i>The</i> | <i>room</i> | <i>was</i> | <i>clean</i> | <i>and</i> | <i>very</i> | <i>nicely</i> | <i>decorated</i> | <i>.</i> |
|------------|-------------|------------|--------------|------------|-------------|---------------|------------------|----------|

Heuristics to handle exceptions

- Separate punctuation marks and quotes from words
decorate. → decorate .
"new" → " new "
- Separate possessive endings or abbreviated forms from preceding words:
Selina's → Selina 's
Hebe's → Hebe is
Ella's → Ella has

2) Entity Detection

Entity Dictionary is predefined which consist of the feature like Outdoor, Activities and Food.

Categorization Rule : Featured keyword matching

Detect entities in sentence → Match Entities in dictionary

If match, the sentence will be assigned to the feature

Related keywords of the features

| Feature | Related Keywords |
|------------------|---|
| Outdoor | outdoor, pool, beach, garden |
| Activities | activity, massage, gym, cycling, fitness center, hot tube, wellness center, spa, children playground, bicycle rental, karaoke |
| Food and Drink | food, drink, bar, coffee house, lounge, buffet, restaurant, breakfast, dinner, chef, waiter, lunch, dishes, all food and drinks menu |
| Internet | Wi-Fi, WLAN, wired internet, internet, Wi-Fi connection, online |
| Parking | parking, car, parking space, parking fee, parking reservation |
| Customer service | Customer service, service, room service, front desk, tour desk, luggage storage, babysitting services, laundry, dry cleaning, meeting/banquet, fax/photocopying, currency exchange, check in, check out, staff , transport, lobby |
| Room | room, room reservation, pillow, single bed, double beds, king-size beds, beds, shower room, shower, bathroom, balcony, air condition, toiletries |
| Environment | environment, hotel, atmosphere, noisy, convenient, location, sea view, balcony, welcomed, view, decoration, design, stay, near, tourists, facilities |

Outdoors(2)



5.00

Parking(1)



5.00

Room(5)



4.00

Environment(12)



3.75

Customer Service(11)



1.82

Food & Drink(4)



1.25

Internet(1)



0.00



Eastern and Oriental

10, Lebuhr Farquhar, 10200 Georgetown, Pulau Pinang

Environment

- We do not feel the exclusive of the welcoming of your hotel staff.
- The hotel is too noisy as rooms are not soundproof.
- Too many clients are very impolite (tourists from China).
- Great hotel steeped in history.
- I like the name of the Hotel.
- The sea behind the hotel is beautiful!
- I like very much the colonial style and location of this Hotel.
- Small little balcony to enjoy the sound of the wave.
- The old world charm of some of the old parts of the hotel.
- Feel very welcome right from the start.
- The environment of the hotel is nice!
- I like the pool and surroundings and the decoration.

3) Post-processing

- Sentiment is extracted from sentences
- Referring to sentiWordNet

Value of the term can be determined

Sentence : The breakfast is tasty.

Entity : breakfast

Sentiment term: tasty

Based on SentiWordNet, information below is provided.

| <i>Term</i> | <i>Description</i> | <i>Positivity</i> | <i>Objectivity</i> | <i>Negativity</i> |
|--------------|--|-------------------|--------------------|-------------------|
| <i>Tasty</i> | <i>pleasing to the sense of taste;</i> | <i>0.625</i> | <i>0.125</i> | <i>0.25</i> |

Based on S.A.Algorithm : (identifying sequence is important)

terms with positivity value ≥ 0.5 , the sentiment values is "1"

terms with objectivity value ≥ 0.5 , the sentiment values is "0.5"

terms with negativity value ≥ 0.5 , the sentiment values is "0"

Thus, "tasty" is defined as a positive term and a value of '1' is assigned to it

Rating

Overall rating of the features will be determined

| Entity | Review | Senti_value | Rating |
|-----------|--|-------------|--------|
| Breakfast | The breakfast offered by this hotel is <u>nice</u> . | 1 | } 4.17 |
| | The breakfast is <u>tasteless</u> . | 0.5 | |
| | Breakfast served is <u>great</u> and variety of choices. | 1 | |

$$\textit{Entity Rating} = \frac{\text{sum of (sentiment value * rating scale)}}{\text{total number of entity related sentences}}$$

Outdoors(2)



5.00

Parking(1)



5.00

Room(5)



4.00

Environment(12)



3.75

Customer Service(11)



1.82

Food & Drink(4)



1.25

Internet(1)



0.00



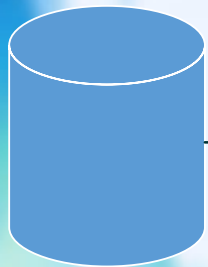
Eastern and Oriental

10, Lebuhraya Farquhar, 10200 Georgetown, Pulau Pinang

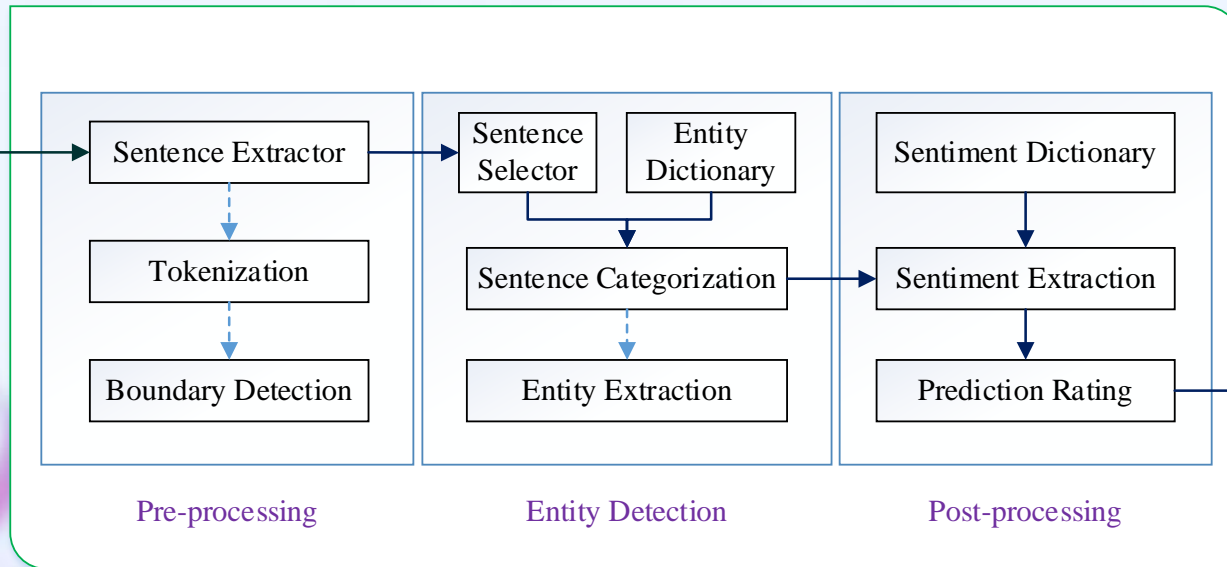
Room

- The hotel is too noisy as rooms are not soundproof.
- beautiful rooms at the victory annex suites.
- I love the interior design of the room especially axenne wing.
- Excellent room design.
- Small little balcony to enjoy the sound of the wave.

Proposed Framework



MySQL Database



Browser

Limitation

Review case 1: Nasi Lemak sambal sucks sour! please learn new recipe from best sambal udang or sambal sotong gerai and learn formmakcik which not stingy to share their recipe. Hard Rock hotel such a prestige and world class pls do something with the local delicacy food and embrace it to be one of the nice local food to eat. Please treat local food as good as international food :).

Unable to provide the proper or accurate data processing when weak grammar or wrong sentence structures is processed

There are some special cases whereby sentiment analysis might not perform well

Review case 2: The car park is full.

Ambiguous meaning

Review case 3 : The interior design of the room is not bad.

The sentiment shifter inside the sentence will change the sentiment orientation.



Conclusion

- Improve readability of reviews
- Save reader time
- Prevent biased user rating